

Human Vulnerability and AI: The Risk of Risk Regulation



Guido Smorto

Abstract The new EU Regulation 2024/1689 on Artificial Intelligence has been subject to many criticisms, especially from those who believe it is not fully adequate to tackle the needs arising from the growing use of AI systems: a limited number of prohibitions, basic transparency obligations for AI systems that present specific risks of impersonation or deception, a risk management system based on documents and conducted by the same provider of the system. Less attention has been paid to the preliminary option made by the EU legislature to adopt a so-called risk-based approach. This article aims to focus on the choice to regulate AI systems through so-called risk regulation and also ponder the consequences of this choice on the distribution of costs and benefits originating from the use of AI systems and the protection of individual rights. In its final part, the article formulates some proposals that support an interpretation of the Regulation more oriented toward risk distribution and protection of rights.

Keywords Artificial Intelligence · EU Artificial Intelligence Act · Risk · Distribution · Liability

1 The European Regulation on Artificial Intelligence

After a long and complex negotiation, in 2024, the European Union finally adopted the European Regulation on Artificial Intelligence (henceforth, Regulation).¹ The leading representatives of the European institutions agreed to call this result a historical one, stressing how this Regulation is a true “first” across the world: the first law

G. Smorto (✉)
University of Palermo, Palermo, Italy
e-mail: guido.smorto@unipa.it

¹ Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

ever on artificial intelligence meant to become a global standard for the development of reliable artificial intelligence, which guarantees security and fundamental rights.

To achieve these ambitious goals, the Regulation adopts a risk-based approach. It categorises AI systems according to whether they generate unacceptable, high, specific, limited, or minimal risks.

To provide rules that are “proportionate” to the interests to be protected and the associated risks, for each of these types of risks it sets prohibitions, mandatory minimum requirements under which to design and develop AI systems, transparency obligations, or simple measures to promote self-regulation.²

Within this typological categorisation, at the heart of the Regulation are “high-risk” AI systems, i.e., those that “have a significant harmful impact on the health, safety and fundamental rights of persons”.³ The definition of an AI system as “high-risk” entails the provider’s obligation to establish, implement, document and maintain a quality management system, i.e., a set of policies, procedures and instructions for the development, control and quality assurance of AI systems, to be applied at the various life stages of the systems in question: design, development, marketing, use, and post-market monitoring.⁴

The risk management system lies at the heart of the quality management system.⁵ In detail, the Regulation requires the provider to adopt a risk management system “proportionate” to the size of its organization,⁶ to draw up the EU declaration of conformity,⁷ to affix the CE marking to the system,⁸ and to implement post-market monitoring to evaluate the continuous compliance of AI systems with the requirements set out in the Regulation.⁹ Wherever an AI system meets these requirements, or namely in other words, where risk mitigation measures are “appropriate” under

² In its final version, the Regulation (EU) 2024/1689, includes general-purpose AI models, with obligations modelled on those set for high-risk AI systems. Cf. Chapter 5.

³ Regulation (EU) 2024/1689, Rec. 46 (“AI systems identified as high-risk should be limited to those that have a significant harmful impact on the health, safety and fundamental rights of persons in the Union and such limitation should minimise any potential restriction to international trade.”); art. 6 (Classification rules for high-risk AI systems) considers an AI system to be high-risk if it is intended to be used as a safety component of a product or if the AI system is itself a product covered by the Union harmonisation legislation. In addition, an AI system is deemed as high-risk if employed in those sectors listed in Annex III (High-risk AI systems referred to in Article 6(2)): biometrics; critical infrastructure; educational and vocational training; employment, workers management and access to self-employment; access to and enjoyment of essential private services and essential public services and benefits; law enforcement; Migration, asylum and border control management.

⁴ Regulation (EU) 2024/1689, Article 17.

⁵ Regulation (EU) 2024/1689, Article 9.

⁶ Regulation (EU) 2024/1689, Article 17, para. 2: “The implementation of the aspects referred to in paragraph 1 shall be proportionate to the size of the provider’s organisation.”

⁷ Regulation (EU) 2024/1689, Article 47.

⁸ Regulation (EU) 2024/1689, Article 48.

⁹ Regulation (EU) 2024/1689, Article 72.

the Regulation, the relevant residual risk associated with each hazard and the overall residual risk of the high-risk AI systems is deemed “acceptable.”¹⁰

Following the approach known as the “New Legislative Framework”,¹¹ the Regulation gives the so-called European standardisation organisations the task of developing - at the request of the Commission - special “harmonised standards”, which translate legal requirements prescribed by the Regulation into detailed technical rules, setting the conditions for the marketing and use of AI systems. High-risk AI systems that comply with such harmonised standards enjoy a presumption of conformity with the Regulation.¹²

2 Critical Reflections on the Regulation

The Regulation has been the subject of much criticism, especially from those who believe it is not fully adequate to handle the dangers associated with the use of artificial intelligence.¹³

With regard to those AI practices that are prohibited as they generate unacceptable risks, the list of prohibitions has been criticised for being extremely limited, due to the combination of different provisions that, on the one hand, establish in detail what is prohibited and, on the other hand, are constructed around a series of explicit or sometimes implicit exceptions that end up broadening the permitted uses of artificial intelligence, including those that are potentially at odds with European values.¹⁴ Regarding the risks of impersonation or deception of those who interact with the AI

¹⁰ Regulation (EU) 2024/1689, Art. 9, par. 5 (“The risk management measures referred to in paragraph 2, point (d), shall be such that the relevant residual risk associated with each hazard, as well as the overall residual risk of the high-risk AI systems is judged to be acceptable.”). Neither the text of the Regulation nor the Explanatory Memorandum offer clarification as to what an acceptable residual risk is. Clause 3.8 of Safety Aspects - Guidelines for Their Inclusion in Standards defines it as “any risk remaining after risk reduction measures have been implemented”; and “acceptable” the level of risk “that is accepted in a given context based on the current values of society” (Clause 3.15). In literature, residual risk is commonly deemed to be the risk which remains after appropriate risk management measures have been implemented, and can be expressed by the formula: residual risk equals risk proper to a certain activity minus risk control. On this point, see, e.g., Tamboli A (2019, p. 103) (Residual Risk = Inherent Risk – Risk Control). Cost-benefit analysis is a central element of risk acceptability assessment. On this point, see Schuett J (2023, p. 13), Fraser and Bello Y Villarino (2022, p. 5): “We find that the greater the emphasis on safety and precaution, the more the cost of residual risks is borne by AI providers (and the higher the barriers to market entry for high-risk AI systems); the greater the emphasis on the social utility of the system, the more likely it is that individuals will carry the residual risks of AI systems –jeopardizing their rights and safety”.

¹¹ European Council (2008) New Legislative Framework adopted on July 9, 2008, as referenced in Annex II of Regulation (EU) 2024/1689.

¹² Regulation (EU) 2024/1689, Article 40.

¹³ Veale and Zuiderveen Borgesius (2021, p. 97), Malgieri and Pasquale (2022, p. 1), Kaminski (2019, p. 1529).

¹⁴ Regulation (EU) 2024/1689, Article 5. For a review, see Ruschemeier (2023, p. 372), Veale and Zuiderveen Borgesius (2021, p. 98).

systems, the absence of substantive limitations on using such systems has also been criticised, as the Regulation prescribes mere transparency requirements.¹⁵

Yet, the most significant reproach is focused on high-risk AI systems. The Regulation has been criticised for using relatively mild risk regulation techniques.¹⁶ It does not provide for a system of licences, time permits, and “resilience” tools, which both literature and practice have identified as characteristic of this regulation.¹⁷ It only resorts to strong precautionary measures - emergency protocols, risk thresholds whose attainment automatically results in the system’s withdrawal from the market, or conditional licences – to a limited extent. It forgoes applying the precautionary principle in favour of a cost–benefit analysis.¹⁸ It places a conformity assessment at the centre, primarily made up of duties to comply with documentary and monitoring obligations. The conformity assessment procedure and the confirmation of compliance with risk mitigation measures must be conducted by the very same providers rather than by third-party, independent bodies.¹⁹ In employing proportionality between risks and benefits as its main principle, it resorts to open standards – adequacy, acceptability, appropriateness – which are left to the interpretation of those directly involved. It chooses a governance model centred on European standardisation organisations.

While critical comments on the Regulation have focused mainly on the limitations of the legal measures adopted for risk mitigation purposes, far less attention has been paid to the preliminary choice of looking at the uses of artificial intelligence through the technical toolkit of risk regulation and to the consequences that such a choice entails.²⁰

¹⁵ Regulation (EU) 2024/1689, Article 50.

¹⁶ Kaminski (2023b, p.1381) defines the system adopted by the Regulation as “ultra-lite licensing”.

¹⁷ Kaminski (2023b, p. 1370), Marchant and Stevens (2017, p. 236 ff).

¹⁸ On the different parameters used in risk assessment, see Coglianese (2020) 1, at 9: “In making these value judgments about risk, decision-makers can adhere to one of four basic normative principles or approaches: (1) Eliminate all risk – or at least all man-made risk (the zero-risk approach); (2) Reduce risk to an acceptable level (the acceptable risk approach); (3) Reduce risk until the cost of doing so reaches an unacceptable level (the feasibility approach); and (4) Balance the benefits of risk reduction with the costs (the efficiency approach)”. On the contrast between the precautionary principle and cost-benefit analysis in risk regulation, see Posner (2004, p. 140), Sunstein (2005); Sunstein (2000, p. 1059). In a comparative perspective, for a comparison between the North American and European models, see Marchant and Stevens (2017, p. 240 ff.), Kysar (2006, p. 1), Wiener (2003, p. 207), who affirms that the conventional between Europe (in favour of the precautionary principle) and the United States of America (which instead adopts cost-benefit analysis) in many cases amounts to an excess of simplification.

¹⁹ The Regulation departs from other harmonised rules that confer the power of conformity assessments on independent authorities, such as the pharmaceutical sector, where the European Medicines Agency (EMA) approves the introduction of new products onto the market and monitors their safety throughout their life span in order to protect public health and ensure the availability of high quality, safe and effective medicines. The point is well argued by Veale and Zuiderveen Borgesius (2021, p. 102). Under the Regulation, only limited biometric systems undergo a conformity assessment by so-called “notified bodies.”

²⁰ De Gregorio and Dunn (2022, p. 473), Chamberlain (2023, p. 1), Schuett (2023, p. 1).

Against this backdrop, the purpose of this article is to focus on the choice that informs the overall logic of the entire regulatory framework of the Regulation and the detailed rules that comprise it: namely, that of regulating AI systems from an ex-ante perspective, based on the model of EU product safety regimes. To this end, after assessing the potential benefits of risk regulation in an area such as artificial intelligence, the article ponders the consequences of adopting such a regulatory model in this field, in order to focus attention on the implications of this choice on the distribution of costs and benefits resulting from the use of AI systems, and also the consequences in terms of protection of fundamental rights and the respect for EU values. In the final part, it formulates some initial interpretive proposals aimed at enhancing the individualised and contextual dimension of the risks generated by artificial intelligence under the Regulation.

3 The Reason Behind the Use of Risk Regulation for Artificial Intelligence

The relatively minor attention paid to the fundamental choice made by the European legislature to adopt risk regulation in the field of AI systems is probably due to the confluence of several different factors. On the one hand, the option chosen by the EU legislature to regulate potential harms from AI systems as risks to be managed is entirely in line with what is happening in the rest of the world. We are, in fact, witnessing a remarkable convergence at the global level of the use of risk regulation as the most appropriate way to govern artificial intelligence, complemented by studies that emphasise the many advantages such an approach would provide in contexts such as those related to the regulation of new technologies and, in particular, artificial intelligence.²¹

Firstly, based on an ex-ante, prognostic, and future-oriented view, risk regulation is deemed particularly well suited to establish flexible rules for a constantly changing field with blurred boundaries. On the one hand, risk regulation would allow for adequate protection of those values and interests that only sometimes translate into vested rights recognised by the legal system, as may often be the case in the digital realm. In a similar vein, this regulatory technique would prove to be particularly suitable for governing potentially harmful consequences that are difficult to predict and quantify accurately, such as those arising from dangers that are not fully known (e.g., those of so-called virtual harms).²²

Similar reasoning applies to the ascertainment of the occurrence of damage, causation and identification of the liable party: something which could prove particularly

²¹ See, e.g., the proposals developed by the National Institute of Standards and Technology (2023) in the USA. Cf AI Risk Management Framework (AI RMF 1.0), p. 13. In literature, see, among others: Guihot et al. (2017, p. 446), Marchant and Stevens (2017, p. 236), Scherer (2016, p. 356), Solow-Niederman (2020, p. 691).

²² Wolfendale (2007, p. 111), Young (2023, p. 1), Carter and Palermos (2016, p. 542).

complex in the event of the development and use of increasingly sophisticated forms of artificial intelligence, also given the opacity of AI systems that is further intensified by the protection of industrial secrets of these systems, famously termed as black boxes.²³ In this context, the decision to adopt the New Legislative Framework, with the resulting involvement of specialised technical bodies in the definition of risk mitigation measures, is also considered the most appropriate for addressing issues of considerable technical complexity, such as those related to the use of artificial intelligence.

4 Civil Liability and Risk Regulation

The line of reasoning that highlights the many benefits of ex-ante regulation is often conducted in a comparison, whether implicit or explicit, with ex-post remedies and, in particular, with torts and insurance law. Drawing on studies in the economic analysis of law,²⁴ it is commonly believed that ex-ante regulation is preferable to a system based on ex-post remedies and compensation in contexts such as the one under consideration.

Following a similar path to that taken in the context of environmental damage,²⁵ many affirm that the technical complexity of AI systems may amount to an insurmountable obstacle to individual access to ex-post forms of compensation, significantly raising the costs that the injured party wishing to take legal action would have to bear, and thus discouraging them from taking any course of action.²⁶ This risk would be even more significant in legal systems that do not offer full recognition of class actions and leave the compensatory remedies to individual plaintiffs.²⁷

The same holds true for damages that are negligible at the individual level but significant at the aggregate level. In such cases, the incentive for individual action is inevitably diluted, and the tort system does not result in an optimal level of deterrence for reasons of collective action.

In light of these limitations, the tort system would end up being merely conjectural and, as a result, would not bring about the deterrent effect and internalisation of risks that typically result from the fear of incurring tortious liability. The many obstacles to a comprehensive and adequate ex-post compensatory type of protection, highlighted in the literature, would produce an overall reduction in tort liability's regulatory (as well as compensatory) function, with the consequence of diluting the deterrent

²³ Pasquale (2015).

²⁴ On the relationship between risk regulation and liability, see, among others, Calabresi (1970), Shavell (1984a, p. 271, 1984b, p. 357), Shavell (1987), Stapleton (1986, 1995, p. 820), Sugarman (1989, p. 555), Schuck (1991, p. 1345), Atiyah (1997); Cane (1997, 2018), Harlow (2004).

²⁵ On the relationship between risk regulation and environmental damages, see Abraham (2002, p. 379), Hylton (2002, p. 515), Schroeder (2002, p. 583), McGarity (2005, p. 371), Abelkop (2014, p. 381), Kysar (2018, p. 48).

²⁶ Froomkin (2015, p. 1713), Hirsch (2006, p. 1), Ben-Shahar (2019, p. 104).

²⁷ Edwards and Veale (2017). In Italian literature, see Comandé (1999).

effect from what is socially desirable.²⁸ It is also from this widespread recognition of the difficulties in the tort system that ex-ante risk-reducing regulation has come to dominate as the preferred model for regulating AI systems.²⁹

5 The Consequences of Risk Regulation

The European discipline on artificial intelligence is often presented by appealing to common sense: the more significant the risk, the stricter the rules, using a basic approach symbolised by the famous pyramid of risks adopted by European institutions to epitomise and disseminate the European approach to artificial intelligence.³⁰

However, far from being a neutral or even obligatory choice, framing the potential harms of AI systems as risks and prescribing measures for their mitigation brings with it a specific technical toolkit and a distinct way of handling the consequences of using artificial intelligence, which amounts to value-laden regulatory decisions.³¹

Firstly, considering the potential harms caused by artificial intelligence as risks to be mitigated, and demanding the adoption of ex-ante technical measures aimed at reducing the risks within acceptable levels implies that the technology in question is permitted despite the harm it may cause.³² As an example, in qualifying as high-risk AI systems, those systems used in several key sectors of society (biometrics; critical infrastructure necessary for the provision of an essential service; education and vocational training; labour; essential private services and essential public services and benefits; law enforcement; migration, asylum and border control; administration of justice and democratic processes),³³ the Regulation permits the use of AI systems in all of these areas, once such systems have appropriate risk mitigation measures in place.³⁴

Secondly, regulating the uses of artificial intelligence according to the tools and the logic of risk regulation implies that problems arising from their use are considered to be such that they can be solved through technical solutions. Risks are identified, their

²⁸ Coglianese (2020, 1).

²⁹ Spindler (1998, p. 316).

³⁰ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

³¹ Kaminski (2023b, p. 1351), Kaminski (2023a, p. 70), Boyd (2012, p. 897) (“As understood here, risk thinking is intended as shorthand for the various concepts, tools, and practices that underwrite the formal understanding and assessment of risk”).

³² Risk is commonly defined as the combination of the probability of the occurrence of an injury and the severity of that injury. Cf. ISO/IEC, Guide 51:2014, Safety aspects – Guidelines for their inclusion in standards. §3.9 defines risk as the combination of the probability of the occurrence of an injury and the severity of that injury. Available at: <https://www.iso.org/standard/53940.html>.

³³ Regulation (EU) 2024/1689, Annex III.

³⁴ Pardolesi and Davola (2020, p. 104), Busch and De Franceschi (2018, p. 408), Casey and Niblett (2017, p. 1401).

probability of occurrence is estimated, a tolerable level of risk is established (“acceptable”, according to the Regulation),³⁵ and appropriate measures to achieve this goal are implemented. The view is “to improve” a specific technology through quantitative solutions rather than critically assessing the suitability of its use in a particular context or field, following a trend described critically as “techno-correctionist”.³⁶ Such an approach, it is noted, fails to consider how artificial intelligence changes the features of the field of activity in which it is employed and, especially, the distribution of costs and benefits.³⁷

Also, in assessing the specific case and identifying appropriate technical solutions, the technical bodies issuing harmonised standards make mathematical and statistical evaluations, adopting an ex-ante prognostic logic, which often encapsulates normative assessments: for example, they may assess the acceptability of the probability of harm to individual rights or groups, or on the appropriate level of risk with which members of a society may be confronted when faced with using a specific AI system. In other words, by employing risk regulation tools, crucial normative judgements are translated into mere technical measures made up of quantitative standards.³⁸

In such cases, risk regulation raises the greatest perplexity. The standardised procedural remedy-based technical and quantitative approach, typical of risk regulation, ends up transforming even non-economic values and interests, from human life to respect for fundamental rights, into monetary standards, making vital choices for society through the use of metrics, technical and documentary obligations and statistical surveys, or rather, through the typical tools of risk regulation.³⁹

Such an approach leaves several interpretive questions open. How can we determine when residual risk is acceptable when fundamental rights are at stake? How do

³⁵ On the difficulties of defining acceptable residual risk, see Fraser and Bello Y Villarino (2022, p. 23).

³⁶ Eaglin (2021, p. 151) (“The challenge, then, is not simply to regulate the tools through law, but to see how law creates the dilemmas that the tools present”). A number of questions arise from this observation: “How has law socially constructed the discourse around technology? How does law facilitate the material conditions that produce the demand for certain technologies? How do technologies reproduce or alter our understanding of race in society, and what is the role of law in perpetuating that meaning?” Hellman (2020, p. 814), Starr (2014, p. 806).

³⁷ Gabriel (2022, p. 218) (“AI is not simply an additional ingredient that supervenes onto a stable practice leaving the fundamental elements of that practice untouched. Rather, AI interacts with the behavior of human decision-makers to shape the character of these practices, including how they distribute benefits and burdens across the population”).

³⁸ Fraser and Bello Y Villarino (2022, p. 12) (“Impact assessment tools, and risk assessment and management standards will help in identifying, quantifying, comparing, prioritising and planning the management of residual risks. But they cannot tell a duty-holder whether in all the circumstances, a risk posed by a specific system deployed in a specific context for a specific purpose is acceptable”).

³⁹ Coglianese (2020, p. 8) (“Put simply, the issue has often been characterized by the question: How safe is safe? Answering this question requires making a normative decision about how much risk members of the public should confront in their lives”), Wagner (1995, p. 1613). With specific reference to AI Regulation see: Schuett (2023, p. 1), Kaminski (2023a), Kim (2017, p. 202), Hellman (2020, p. 834), Ajunwa (2020, p. 1679), Ackerman and Heinzerling (2004), Kelman (1981, p. 33). With specific regard to data, see Barocas and Selbst (2016, p. 671), Kim (2017, p. 202).

we decide, for example, when a risk of discrimination is acceptable if the expenditure of one hundred allows for a fifty per cent reduction in risk, while achieving a seventy per cent reduction requires spending three hundred (assuming that such an assessment is possible)? In such a case, is it acceptable to limit risk reduction to only fifty percent, given the costs to be incurred?⁴⁰ The extensive use of open standards throughout the entire Regulation amplifies the issue, deferring subtle balancing acts to technical bodies that translate them into technical prescriptions.⁴¹ In conclusion, the actuarial and technocratic logic of risk regulation systems encapsulates fundamental policy choices that must be rendered explicit.

Risk regulation is an appropriate regulatory technique for managing easily identifiable and quantifiable risks. However, resorting to standardised technical measures aimed at rationalising and quantifying risks through statistical, actuarial, and insurance tools according to mathematical calculations may not be entirely appropriate for fundamental rights and values that do not lend themselves easily to quantification. In other words, risk regulation is more suitable for estimating quantifiable risks and much less so for those more challenging to domesticate in terms of numbers.

Despite using the language of scientific certainty and hard science, this type of cost–benefit analysis involves making complex normative judgments, such as the ones typically arising from the violation of rights, where it is far more problematic to use technical tools for evaluating and measuring activities.⁴²

The danger is to overlook the risks of violating fundamental rights or key societal values, making these risks less visible. Correspondingly, by framing delicate policy issues as neutral technical choices, the Regulation leaves the definition of broad and indefinite notions – such as fairness, diversity or non-discrimination, among others – to technical standards set by private entities.⁴³

Moreover, the characteristic standardisation of risk regulation, and the resulting formulation of complex normative judgments on how to assess benefits and risks, is even more problematic in an area, such as that of the use of artificial intelligence, where the development of shared ethical standards is far from being reached,⁴⁴ and the risks are not yet entirely known due to the absence of an established case history.⁴⁵ These difficulties are further exacerbated by empirical uncertainties that often affect

⁴⁰ In recalling the proportionality between risks and measures to be taken, under the Regulation, one of the central elements of said proportionality judgment is the costs incurred. See, e.g., Regulation (EU) 2024/1689, Recital 146

⁴¹ Ebers et al. (2021, p. 592), Solow-Niederman (2023, p. 2), Laux et al. (2024).

⁴² Kaminski (2023b, p. 1378) (“This, then, is one of the core challenges for AI risk regulation: it deploys a largely ex ante regulatory tool best suited for readily quantifiable harms to address big, often-unquantifiable, often-contested, often-contextual, and often-individualized ‘risks’”).

⁴³ Veale and Zuiderveen Borgesius (2021).

⁴⁴ Schuett (2023, p. 13).

⁴⁵ Fraser and Bello Y Villarino (2022, p. 12) (“Developing standards is complicated in the context of AI. Standardisation works well in contexts where risks are known and there is good evidence from past events, less well when data from previous experience is missing, and potentially terribly when a system has not been tested in real life or can change its expected behaviour through a process of learning.”). See also Kleindorfer and Kunreuther (1999, p. 149–155), Turner (2019, p. 75).

risk regulation, as testified by the considerable diversity in results obtained from using this regulatory technique in different fields.⁴⁶

6 Risk Distribution

Having considered the many reasons why the seemingly neutrality of risk management through standardised technical norms is contradicted by the need for normative judgments embedded in technical choices, as well as by frequent empirical difficulties, we can now focus on the distributional effects of this normative technique when fundamental rights and values are at stake.

The importance of in-depth reflection on the distribution of risks and benefits associated with artificial intelligence and the problematic issues involved in using risk regulation stem from the fact that the potential risks and benefits resulting from the use of a particular AI system rarely fall on the same individuals or groups.

As made explicit in the recitals of the Regulation, society takes certain risks in the name of the benefits that a specific choice may bring. This assessment is performed in aggregate, looking at the overall balance between costs and benefits, leaving considerations of whether these are distributed to different degrees over distinct segments of the population in the background.

In the aggregate, adopting technologies that carry high risks for individuals or groups is perfectly rational if more significant societal benefits offset these risks. Indeed, the standardisation of ex-ante risk regulation systems does not consider the individualised and contextual nature of some types of risks and the biases that AI systems may bring.⁴⁷ The use of artificial intelligence, for example, to manage state welfare, select job applicants or grant a mortgage, can have overall positive effects for the state, the employer or the bank that adopts such systems, with undeniable benefits, for example, in terms of efficiency. At the same time, adopting these AI systems may produce detrimental effects and risks of discrimination for specific individuals or groups who are particularly disadvantaged by such systems.⁴⁸

In addition, treating potential harms as risks to be managed not only means adopting an aggregate view that looks at the overall impact of a given risk on society, to the detriment of specific individuals or groups. In most cases, the most vulnerable individuals or groups are those most exposed to the detrimental consequences of choices based on the overall impact of these systems.⁴⁹ This may occur mainly due to the fact that the standards and thresholds used to make this risk assessment often

⁴⁶ Coglianesi (2020, p. 12) (“Currently, the estimated net benefits of risk regulation vary markedly across, and within, policy areas. Some regulations have been estimated to cost only in the tens of thousands of dollars per life saved; others have been estimated to cost billions of dollars for each life saved”). See also Breyer (1993).

⁴⁷ Chamberlain (2023, p. 7).

⁴⁸ Eubanks (2018), Umoja Noble (2018), Benjamin (2019).

⁴⁹ Barocas and Selbst (2016, p. 684).

refer to an average standard, such as the archetypical “reasonable person”, to the detriment of a more nuanced and subtle view of society, which takes into account the differences and diversities between individuals and between groups.⁵⁰

7 Risk Regulation and Ex-Post Liability

The flaws of a strategy solely based on risk mitigation are further exacerbated by the fact that the Regulation does not expressly provide for ex-post compensatory remedies (thus, marking a departure from choices made in the past, as in the case of the GDPR). The risk is to leave those who suffer the harm without remedies or with insufficient remedies, undermining the possibility of obtaining protection once the harm has occurred.

With specific regard to the Regulation, the crucial question is the compensation of so-called “residual risk”, i.e., the risk that remains even after appropriate mitigation measures have been taken; namely, whether it is possible to get compensation in the case of AI systems that have an “acceptable” level of risk, based on the ex-ante assessment made in compliance with the Regulation.

In this regard, a preliminary aspect of the relationship between ex-ante risk regulation and ex-post liability must be clarified. In principle, the Regulation merely dictates the conditions under which AI systems can be placed on the market. At the same time, the consequences arising from using AI systems, once placed on the market, are still subject to the ordinary rules governing the area they operate in. For example, the actions of an employer who uses an AI system to fire, promote or hire a person are, per se, susceptible to judicial review in the same way that decisions made by the same employer without using AI systems would be.

In principle, we can limit our analysis to recognising the formal autonomy of the two systems – ex-ante risk regulation and ex-post tortious liability – and conclude for the autonomy of the regulation of the conditions for placing AI systems on the market and the liability regimes for damages resulting from the use of these systems. However, such a conclusion neglects to evaluate the overall effects of ex-ante regulation on the enforceability of ex-post remedies to protect individual rights. While not excluding liability for damages arising from AI systems, the peril is that compliance with legally prescribed risk mitigation measures may, in many cases, result in a considerable obstacle to the affirmation of individual, ex-post liability.

Such a danger is indisputably more severe in the case of fault-based liability. The many similarities between risk assessment criteria and the factors commonly considered for assessing fault in tortious liability reveal this potential overlap. The similarity of the elements involved in risk management and liability regimes heightens the possibility that compliance with technical standards may be considered sufficient

⁵⁰ Boyd (2012, pp. 912 and 927).

to exclude fault, with an exchange of doctrines and evaluations that the role of expert witnesses in litigation may foster.⁵¹

In principle, more room should still be allowed for compensatory damages under a strict liability regime. However, even in this case, it is easy to imagine numerous obstacles to compensation for damages arising from AI systems that comply with the Regulation.⁵² What makes individual liability marginal, even in the case of strict liability, is not only the limited scope of the strict liability regime in most Member States. Also, many doctrines applicable to strict liability may end up significantly limiting compensatory claims related to AI systems, even when a strict liability regime applies: as an example, for risks that cannot be reasonably mitigated or eliminated through the development or design of the AI system or by providing adequate technical information (so-called, unavoidably unsafe products doctrine)⁵³; or for those unknown risk and harms, in light of the state of scientific and technical knowledge.⁵⁴

In the past, the European Parliament suggested extending the area of strict liability to damages caused by high-risk AI systems and applying a special fault-based liability regime to other AI systems, also providing for a presumption of fault.⁵⁵ In this way,

⁵¹ Cafaggi (2006, p. 199) (“To the extent that regulation has increasingly provided new standards for examining the conduct of potential injurers, the corresponding modifications in civil liability have proved significant for the definition of standards in negligence and strict liability”); Kysar (2018, p. 51).

⁵² Chamberlain (2023, p. 8) draws a parallel between the regulation of high-risk AI systems and the imposition of strict liability regimes.

⁵³ Cf. art. 9, para. 3 (The risks referred to in this Article shall concern only those which may be reasonably mitigated or eliminated through the development or design of the high-risk AI system, or the provision of adequate technical information.”). In literature, see, e.g., Page (1996, p. 87).

⁵⁴ See Recital 64 AI Act (“The measures adopted by the providers to comply with the mandatory requirements of this Regulation should take into account the generally acknowledged state of the art on AI”); Recital 121 (“Standardisation should play a key role to provide technical solutions to providers to ensure compliance with this Regulation, in line with the state of the art”); art. 8 (“High-risk AI systems shall comply with the requirements laid down in this Section, taking into account their intended purpose as well as the generally acknowledged state of the art on AI and AI-related technologies.”); art. 25 para. 4 (“The provider of a high-risk AI system and the third party that supplies an AI system, tools, services, components, or processes that are used or integrated in a high-risk AI system shall, by written agreement, specify the necessary information, capabilities, technical access and other assistance based on the generally acknowledged state of the art”); art. 50 para. 2 (“Providers shall ensure their technical solutions are effective, interoperable, robust and reliable as far as this is technically feasible, taking into account the specificities and limitations of various types of content, the costs of implementation and the generally acknowledged state of the art, as may be reflected in relevant technical standards.”); art. 51 para. 3 (“The Commission shall adopt delegated acts in accordance with Article 97 to amend the thresholds listed in paragraphs 1 and 2 of this Article, as well as to supplement benchmarks and indicators in light of evolving technological developments); art. 96 para. 1 (“The guidelines referred to in the first subparagraph of this paragraph shall take due account of the generally acknowledged state of the art on AI”).

⁵⁵ European Parliament Resolution (2020) On a liability regime for artificial intelligence (2020/2014(INL)), OJ C 404/05 (2021). In 2022, the EU Commission had also submitted a Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive), COM(2022) 496 final, which was not followed up.

high-risk AI systems would have been equated to dangerous activities, such as those that typically warrant the application of a strict liability regime. However, the proposal remained a dead letter, and, in any case, would have charted a course that would have been difficult to follow, given the diversity - within the European Union - of areas subject to strict liability in national legal regimes and the considerable difficulties of applying a special liability regime to a category - that of high-risk AI systems - that has blurred boundaries.

8 Concluding Remarks

So far, we have highlighted the various reasons why a system of risk regulation adopts an overall, aggregated and standardised view of the risks arising from the use of AI systems, at the expense of an assessment more focused on the individualised and contextual dimension of the consequences stemming from the use of these systems, which pays closer attention to the individuals and groups on whom these risks fall. Moreover, we have pointed out how this disregard for risk distribution and protection of individual rights mainly affects the most vulnerable.

In the conclusion, we provide some interpretive suggestions aimed at fostering an understanding of the Regulation that is more oriented toward distributional considerations and leaves more space for those needs that the Regulation recognises to a limited extent only.

To complement the aggregate and standardised assessment of risks with a consideration of the distribution of these risks, a first group of provisions concerns how the powers reserved by the Regulation for the EU Commission are to be exercised. The first is the procedure the Regulation prescribes for qualifying AI systems as high-risk under art. 7 of the Regulation, which confers the power to identify these systems upon the Commission. In exercising its power to define AI systems as high-risk, the Commission should not only consider the potential extent of harm or negative impact in terms of intensity and capacity to affect more than one person but also give special consideration to the capacity to disproportionately affect a particular group, a feature added only in the latest version of Regulation, which was then adopted.⁵⁶

In a complementary manner, the common specifications, adopted by the Commission in all cases where harmonised standards do not sufficiently consider fundamental rights concerns,⁵⁷ should play a central role, despite the express reference to their exceptional nature in the Regulation.⁵⁸

⁵⁶ Regulation (EU) 2024/1689, art. 7, para. 2(f) (“When assessing the condition under paragraph 1, point (b), the Commission shall take into account the following criteria: (...) ‘the potential extent of such harm or such adverse impact, in particular in terms of its intensity and its ability to affect multiple persons or to disproportionately affect a particular group of persons’”).

⁵⁷ Regulation (EU) 2024/1689, art. 41.

⁵⁸ Regulation (EU) 2024/1689, Recital 121 defines common specifications as an “exceptional” solution (“The common specification should be an exceptional fall back solution to facilitate the provider’s obligation to comply with the requirements of this Regulation, when the standardisation

Also, in formulating requests to draft a harmonised standard to the relevant European organisations under Article 40, the Commission should emphasise those normative questions that stem from developing those technical solutions that European organisations are called upon to lay down.

Finally, when establishing a post-market monitoring plan under article 72(3), and in determining the obligations of providers related to post-market monitoring of an AI system, the Commission should consider any unfavourable effects of an AI system on specific individuals or segments of the population, and take this assessment into account when determining appropriate corrective measures, especially where the most vulnerable are affected.⁵⁹

As is evident, the way the EU Commission will exercise the prerogatives granted by the Regulation will be a decisive factor in giving greater or lesser relevance to distributional considerations in assessing the impact of AI systems. At the same time, other norms of the Regulation are of the utmost importance for our analysis.

The fundamental rights impact assessment for high-risk AI systems also plays a vital role.⁶⁰ While not applicable across the board, but only to “deployers that are bodies governed by public law, or are private entities providing public services”, as well as to specific high-risk AI systems, the individual and contextual dimensions of the impact of AI may be attained by putting a particular emphasis on some of those aspects listed in art. 27: primarily, the categories of natural persons and groups likely to be affected by its use and the specific risks of harm likely to impact the categories of natural persons or groups of persons.⁶¹

Lastly, the many limitations underscored in these pages are further exacerbated by the possible obstacles to ex-post liability regimes. Hence, it is crucial to avoid the Regulation from weakening liability regimes for AI systems, as this would undermine the effectiveness of risk regulation by depriving it of ex-post feedback that may complement, rather than replace, ex-ante risk determination and mitigation. This is even more pressing where risk homogeneity is low, i.e., where the differentiation in the degree of risk is high due to the different actors involved. When this is the case, i.e. whenever the risk is highly contextual and individualised, the ex-post liability system may be more efficient than an ex-ante system of risk regulation.⁶² In such a context, civil liability may play a critical role, not only as a way to provide compensation but also as an incentive for an efficient level of activities, thus taking a strategic

request has not been accepted by any of the European standardisation organisations, or when the relevant harmonised standards insufficiently address fundamental rights concerns, or when the harmonised standards do not comply with the request, or when there are delays in the adoption of an appropriate harmonised standard.”).

⁵⁹ Gabriel (2022, p. 225) applies a Rawlsian approach to technology (“When AI is integrated into a key social practice, in a way that affects the overall distribution of benefits and burdens, it is pertinent to ask whether it does the most it possibly can do to improve the position of the least advantaged member of that system.”).

⁶⁰ Regulation (EU) 2024/1689, art. 27.

⁶¹ Regulation (EU) 2024/1689, art. 27, para. 1(c–d).

⁶² Chamberlain (2023, p. 7), Cafaggi (2006, p. 191 ff.), Kysar (2018, p. 52). On the relationship between tort liability and distributive justice, see Gardner (2019).

regulatory function that is imperative in the legal challenge of governing artificial intelligence by protecting the fundamental rights and values of the European Union.

References

- Abelkop ADK (2014) Tort Law as an Environmental Policy Instrument. *Or. L. Rev.* 92(2): 381–469
- Abraham KS (2002) The relation between civil liability and environmental regulation: An analytical overview. *Washburn Law J.* 41: 379–398
- Ackerman F, Heinzerling L (2004) *Priceless: on knowing the price of everything and the value of nothing*. The New Press
- Ajunwa I (2020) The paradox of automation as anti-bias intervention. *Cardozo Law Rev* 41:1671
- Atiyah P (1997) *The damages lottery*. Hart Publishing
- Barocas S, Selbst AD (2016) Big data's disparate impact. *Calif Law Rev* 104:671
- Benjamin R (2019) *Race after technology: Abolitionist tools for the new Jim Code*. Polity, Cambridge
- Ben-Shahar O (2019) Data pollution. *J. Legal Analysis* 11: 104–159
- Boyd W (2012) Genealogies of risk: Searching for safety, 1930s-1970s. *Ecology L.Q.* 39: 895–987
- Breyer SG (1993) *Breaking the vicious circle: Toward effective risk regulation*. Harvard University Press
- Busch C, De Franceschi A (2018) Granular legal norms: Big data and the personalization of private law. In: V. Mak, E. Tjong Tjin Taj, A. Berlee (eds) *Research handbook on data science and law*. Edward Elgar Publishing Limited, pp 408–424
- Cafaggi F (2006) The legitimacy of private regulators vis-à-vis third parties: Rights, duties, and liability. *EUI Working Paper LAW 2006/32*, p 191
- Calabresi G (1970) *The costs of accidents: A legal and economic analysis*. Yale University Press
- Cane P (1997) *The anatomy of tort law*. Hart Publishing
- Cane P (2018) *Atiyah's accidents, compensation and the law*, 9th edn. Cambridge University Press
- Carter JA, Palermos SO (2016) Is having your computer compromised a personal assault? The ethics of extended cognition. *J. Am. Philos. Assoc.* 2(4): 542–560
- Casey AJ, Niblett A (2017) The death of rules and standards. *Indiana Law Journal* 92(4):1401–1447
- Chamberlain J (2023) The risk-based approach of the European Union's proposed artificial intelligence regulation: Some comments from a tort law perspective. *Eur J Risk Reg* 14(1):1–13
- Coglianese C (2020) The law and economics of risk regulation, in *Penn Carey Law: Legal Scholarship Repository*. All Faculty Scholarship. 2157. <https://ssrn.com/abstract=3552703>
- Comandè G (1999) *Risarcimento del danno alla persona e alternative istituzionali*. Giappichelli
- De Gregorio G, Dunn P (2022) The European risk-based approaches: Connecting constitutional dots in the digital age. *Common Market Law Review* 59(2): 473–500
- Eaglin JM (2021) When critical race theory enters the law & technology frame. *Mich. J. Race & L.* 26: 151–168
- Ebers M, Hoch VRS, Rosenkranz F, Ruschemeier H, Steinrötter B (2021) The European commission's proposal for an artificial intelligence act - A critical assessment by members of the Robotics and AI Law Society (RAILS). *J.* 4: 589–603
- Edwards L, Veale M (2017) Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. *Duke Law Technol Rev* 16:18
- Eubanks V (2018) *Automating inequality: How high-tech tools profile, police, and punish the poor*. St Martin's Press, New York
- European Commission (2021) *Regulatory framework on AI*. Available at: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- Fraser H, Bello Y, Villarino JM (2022) Where residual risks reside: a comparative approach to art 9(4) of the European Union's proposed AI Regulation. <https://ssrn.com/abstract=3960461>

- Froomkin AM (2015) Regulating mass surveillance as privacy pollution: Learning from environmental impact statements. *U. Ill. L. Rev.* 2015(5): 1713–1790
- Gabriel I (2022) Toward a theory of justice for artificial intelligence. *Daedalus* 151(2): 218–231
- Gardner J (2019) *Torts and other wrongs*. Oxford University Press
- Guihot M, Matthew AF, Suzor NP (2017) Nudging robots: innovative solutions to regulate artificial intelligence. *Vand. J. Ent. & Tech. L.* 20(2): 385–456
- Harlow C (2004) *State liability, tort and beyond*. Oxford University Press
- Hellman D (2020) Measuring algorithmic fairness. *Va. L. Rev.* 106(4): 811–866
- Hirsch DD (2006) Protecting the inner environment: What privacy regulation can learn from environmental law. *Ga Law Rev* 41:1
- Hylton KN (2002) When should we prefer tort law to environmental regulation? *Washburn Law J.* 41: 515–534
- Kaminski ME (2019) Binary governance: Lessons from the GDPR’s approach to algorithmic accountability. *S. Cal. L. Rev.* 92: 1529–1616
- Kaminski ME (2023a) The developing law of AI: A turn to risk regulation. *The Digital Social Contract: A Lawfare Paper Series*. U of Colorado Law Legal Studies Research Paper No. 24-5, pp 1–26
- Kaminski ME (2023b) Regulating the risks of AI. *B.U. L. Rev.* 103: 1347–1411
- Kelman S (1981) Cost-benefit analysis: an ethical critique. *Reg Policy* 5:33
- Kim PT (2017) Auditing algorithms for discrimination. *U. Pa. L. Rev. Online* 166: 189–203
- Kleindorfer PR, Kunreuther HC (1999) Challenges facing the insurance industry in managing catastrophic risks. In: K.A Froot (ed.) *The financing of catastrophe risk*. University of Chicago Press, pp 149–194
- Kysar DA (2006) It might have been: Risk, precaution and opportunity cost. *J. Land Use & Env’t L.* 22: 1–58
- Kysar DA (2018) The public life of private law: Tort law as a risk regulation mechanism. *Eur. J. Risk Reg.* 9: 48–65
- Laux J, Wachter S, Mittelstadt B (2024) Three pathways for standardisation and ethical disclosure by default under the European Union Artificial Intelligence Act. *Comput Law Secur Rev* 53:105957
- Malgieri G, Pasquale F (2022) From transparency to justification: Toward ex ante accountability for AI. *Brooklyn L. School. Legal Studies Paper No. 712: 1–27* <https://ssrn.com/abstract=4099657>
- Marchant GE, Stevens YA (2017) Resilience: A new tool in the risk governance toolbox for emerging technologies. *U.C. DAVIS L. REV.* 51: 233–271
- McGarity TO (2005) Regulation and litigation: Complementary tools for environmental protection. *Colum. J. Envtl. L.* 30: 371
- Page JA (1996) Liability for unreasonably and unavoidably unsafe products: Does negligence doctrine have a role to play. *Chi.-Kent L. Rev.* 72: 87–128
- Pardolesi R, Davola A (2020) Algorithmic legal decision making: la fine del mondo (del diritto) o il paese delle meraviglie? *Questione Giustizia* 1: 104–111
- Pasquale F (2015) *The Black Box Society: the secret algorithms that control money and information*. Harvard University Press
- Posner R A (2004) *Catastrophe: Risk and response*. Oxford University Press, Oxford
- Ruscheimer H (2023) AI as a challenge for legal regulation. The scope of application of the Artificial Intelligence Act proposal. *ERA Forum* 23: 361–376
- Scherer MU (2016) Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. *Harv J Law Technol* 29:353
- Schroeder CH (2002) Lost in the translation: What environmental regulation does that tort cannot duplicate. *Washburn Law J.* 41: 583–606
- Schuck P (1991) *Tort law and the public interest: Competition, innovation, and consumer welfare*. Norton
- Schuett J (2023) Risk management in the Artificial Intelligence Act. *Eur. J. Risk Reg.* 1–19

- Shavell S (1984a) A model of the optimal use of liability and safety regulation. *RAND J Econ* 15:271–280
- Shavell S (1984b) Liability for harm versus regulation of safety. *J Legal Stud* 13:357–374
- Shavell S (1987) *Economic analysis of accident law*. Harvard University Press
- Solow-Niederman A (2020) Administering Artificial Intelligence. *S. Cal. L. Rev.* 93: 633–695
- Solow-Niederman A (2023) Can AI standards have politics? *UCLA L. Rev. Disc.* 71: 2–17
- Spindler G (1998) Market Processes, Standardization, and Tort Law. *European Law Journal* 4: 316–336
- Stapleton J (1986) *Disease and the compensation debate*. Oxford University Press
- Stapleton J (1995) Tort, insurance and ideology. *Modern Law Review* 60: 820–845
- Starr SB (2014) Evidence-based sentencing and the scientific rationalization of discrimination. *Stanford Law Rev* 66:803–872
- Sugarman D (1989) *Doing away with personal injury law*. Bloomsbury Academic
- Sunstein C R (2000) *Behavioral law and economics*. Cambridge University Press
- Sunstein C R (2005) *Laws of fear: Beyond the precautionary principle*. Cambridge University Press
- Tamboli A (2019) *Keeping Your AI Under Control: A Pragmatic Guide to Identifying, Evaluating, and Quantifying Risks*. Apress Berkeley, New York
- Turner J (2019) *Robot rules: Regulating artificial intelligence*. Palgrave Macmillan
- Umoja Noble S (2018) *Algorithms of oppression: How search engines reinforce racism*. NYU Press, New York
- Veale M, Zuiderveen Borgesius F (2021) Demystifying the draft EU Artificial Intelligence Act – analysing the good, the bad, and the unclear elements of the proposed approach. *Comput Law Rev Int* 4:97–112
- Wagner WE (1995) The science charade in toxic risk regulation. *Columbia Law Rev* 95:1613–1723
- Wiener J B (2003) Whose Precaution After All? A Comment on the Comparison and Evolution of Risk Regulatory System. *Duke J. Compar. & Int'l L.* 13: 207–262
- Wolfendale J (2007) My avatar, my self: Virtual harm and attachment. *Ethics Inf Technol* 9:111–119
- Young G (2023) What would strong AI understand consent to mean, and what are the implications for sexbot rape? *AI and Ethics* 5:579–590

Guido Smorto, Full Professor of Comparative Law at the University of Palermo and EU Jean Monnet Chair in “Comparative and European Digital Law.” Expert advisor for the EU Commission and Parliament on digital law. Co-author of *Comparative Law. A Very Short Introduction* (OUP, 2023) and *Diritto privato dei mercati digitali* (Mondadori, 2024, 2nd ed.).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

